



별첨 사본은 아래 출원의 원본과 동일함을 증명함.

This is to certify that the following application annexed hereto is a true copy from the records of the Korean Intellectual Property Office.

출원 번호 : 10-2002-0067298
Application Number PATENT-2002-0067298

출원 년 월 일 : 2002년 10월 31일
Date of Application OCT 31, 2002

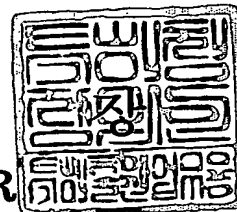
출원 인 : 주식회사 바이오인프라
Applicant(s) BIOLNFRA INC.



2002 년 12 월 06 일

특 허 청

COMMISSIONER



PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

【서지사항】

【서류명】	특허출원서
【권리구분】	특허
【수신처】	특허청장
【제출일자】	2002.10.31
【발명의 명칭】	프로테옴 이미지 마이닝을 이용한 암 질환 분석 방법 및 시스템
【발명의 영문명칭】	Method and System for Analysis of Cancer Biomarkers Using Proteome Image Mining
【출원인】	
【명칭】	주식회사 바이오인프라
【출원인코드】	1-2001-038518-0
【대리인】	
【성명】	손민
【대리인코드】	9-1999-000420-6
【포괄위임등록번호】	2001-055511-9
【대리인】	
【성명】	이세진
【대리인코드】	9-2000-000320-8
【포괄위임등록번호】	2001-055509-9
【대리인】	
【성명】	김성남
【대리인코드】	9-1998-000150-9
【포괄위임등록번호】	2001-055510-1
【발명자】	
【성명의 국문표기】	김철우
【성명의 영문표기】	KIM, Chul Woo
【주민등록번호】	520215-1063227
【우편번호】	471-080
【주소】	경기도 구리시 갈매동 산2-120
【국적】	KR
【발명자】	
【성명의 국문표기】	박영미
【성명의 영문표기】	PARK, Young Mee

【주민등록번호】	611001-2067018
【우편번호】	158-050
【주소】	서울특별시 양천구 목동 목동아파트 1303동 907호
【국적】	KR
【발명자】	
【성명의 국문표기】	박종서
【성명의 영문표기】	PARK, Jong Sou
【주민등록번호】	600902-1074416
【우편번호】	412-270
【주소】	경기도 고양시 덕양구 화정동 948번지 기산아파트 919-202
【국적】	KR
【발명자】	
【성명의 국문표기】	지승도
【성명의 영문표기】	CHI, Sung Do
【주민등록번호】	590219-1011810
【우편번호】	411-310
【주소】	경기도 고양시 일산구 일산동 후곡마을 610동 1303호
【국적】	KR
【발명자】	
【성명의 국문표기】	온승엽
【성명의 영문표기】	OHN, Syng Yup
【주민등록번호】	620320-1480213
【우편번호】	140-751
【주소】	서울특별시 용산구 서빙고동 241-21 신동아아파트 6동 1109호
【국적】	KR
【발명자】	
【성명의 국문표기】	황수찬
【성명의 영문표기】	HWANG, Soo Chan
【주민등록번호】	610119-1030028
【우편번호】	411-370
【주소】	경기도 고양시 일산구 주엽동 강선마을 601-803
【국적】	KR

【우선권 주장】**【출원국명】**

KR

【출원종류】

특허

【출원번호】

10-2002-0019028

【출원일자】

2002.04.08

【증명서류】

미첨부

【심사청구】

청구

【취지】

특허법 제42조의 규정에 의한 출원, 특허법 제60조의 규정에 의한 출원심사를 청구합니다. 대리인

손민 (인) 대리인

이세진 (인) 대리인

김성남 (인)

【수수료】**【기본출원료】**

20 면 29,000 원

【가산출원료】

28 면 28,000 원

【우선권주장료】

1 건 26,000 원

【심사청구료】

22 항 813,000 원

【합계】

896,000 원

【감면사유】

소기업 (70%감면)

【감면후 수수료】

287,000 원

【첨부서류】

1. 요약서·명세서(도면)_1통

【요약서】

【요약】

본 발명은 이미지 마이닝 기법에 의해 혈청 프로테옴의 표본을 생성하고 이를 이용하여 암 질환 여부를 확인할 수 있는 암 질환 분석 방법 및 이를 이용한 시스템을 제공한다. 또한 본 발명은 암 진단에 사용하는 암 특이적 진단 마커를 제공한다.

본 발명의 암 질환 분석 방법은 정상적인 개체 및 질환을 가진 개체의 혈청 프로테옴을 2D-Gel을 이용하여 2차원 이미지화한 후, 이를 이미지 마이닝 기법을 적용하여 혈청 프로테옴 표본으로 데이터베이스화하는 과정, 분석 대상 혈청 프로테옴을 입력받아 이미지화한 후 데이터베이스에 저장된 혈청 프로테옴 표본과 비교하여 암 질환 여부를 판단하는 과정을 포함한다.

또한, 본 발명의 암 질환 분석 시스템은 다수의 정상 개체 및 암 질환 개체로부터 채취한 혈청 프로테옴을 입력받아 2차원 이미지로 변환하여 특징을 추출하고, 특징 데이터 중 최적의 특징 데이터를 식별하여 프로테옴 표본을 생성하는 한편, 분석 대상 혈청 프로테옴을 2차원 이미지로 변환하고 특징을 추출하기 위한 프로테옴 표본 생성 수단, 프로테옴 표본 생성 수단에 의해 추출된 분석 대상 혈청 프로테옴의 패턴과 프로테옴 표본의 패턴의 유사 여부를 판단하기 위한 프로테옴 비교 수단, 프로테옴 비교 수단의 비교결과 분석 대상 혈청 프로테옴의 패턴이 정상인의 혈청 프로테옴 패턴과 유사한 경우에는 분석 대상 혈청 프로테옴을 정상으로 판단하고 그렇지 않을 경우에는 암 질환이 발병한 것으로 판단하기 위한 질환 분석 수단을 구비한다.

이와 같은 암 질환 분석 방법 또는 시스템에 의하면 다양한 혈청 프로테옴을 이미지 마이닝 기법을 적용하여 데이터베이스화하고, 분석 대상 혈청 프로테옴을 프로테옴 표본과 비교함으로써 암 질환 발병 여부를 용이하게 확인할 수 있다.

【대표도】

도 1

【색인어】

암, 프로테옴, 서포트 벡터 머신, 이미지 마이닝, 퍼지

【명세서】**【발명의 명칭】**

프로테옴 이미지 마이닝을 이용한 암 질환 분석 방법 및 시스템{Method and System for Analysis of Cancer Biomarkers Using Proteome Image Mining}

【도면의 간단한 설명】

도 1은 본 발명에 의한 암 질환 분석 시스템의 구성도,
도 2는 도 1에 도시한 프로테옴 표본 생성 수단의 상세 구성도,
도 3은 본 발명에 의한 암 질환 분석 방법을 설명하기 위한 흐름도,
도 4는 도 3에 도시한 프로테옴 표본 생성 방법을 설명하기 위한 흐름도,
도 5는 본 발명에 적용되는 2차원 혈청 프로테옴 이미지의 일 예시도,
도 6은 혈청 프로테옴을 입력받아 프로테옴 표본을 생성하는 과정을 도시화한 것,
도 7은 서포트 벡터 머신에 의한 최적 분리면을 설명하기 위한 도면,
도 8은 SVM/GA를 이용한 유방암 진단의 트레이닝 단계를 나타낸 도면,
도 9는 SVM/GA를 이용한 유방암 진단의 테스트 단계를 나타낸 도면,
도 10은 유방암 진단을 위하여 26개 스팟들을 최적의 특징 데이터로 하여 본 발명의 시스템을 적용한 실제 진단 결과를 나타낸 도면, 및
도 11은 유방암 진단을 위하여 48개 스팟들을 최적의 특징 데이터로 하여 본 발명의 시스템을 적용한 실제 진단 결과를 나타낸 도면이다.

<도면의 주요 부분에 대한 부호 설명>

10 : 암 질환 분석 시스템 102 : 프로테옴 표본 생성 수단
104 : 프로테옴 비교 수단 106 : 질환 분석 수단
108 : 입출력 인터페이스 110 : 제어부
112 : 입력 수단 114 : 출력 수단
116 : 데이터베이스 210 : 전처리 수단
212 : 이미지 처리 수단 214 : 특징 추출 수단
220 : 진화적 분류 수단 222 : GA 처리 수단
224 : SVM 적용 수단 230 : 퍼지 룰 기반 분류 수단
232 : 데이터 맵핑 수단 234 : 룰 기반 분류 수단

【발명의 상세한 설명】

【발명의 목적】

【발명이 속하는 기술분야 및 그 분야의 종래기술】

<22> 본 발명은 다수의 정상적인 생체의 혈청 단백질 및 질병을 가진 생체의 혈청 단백질의 2D-Gel 분리 상태를 이미지로 변환하여, 이미지 마이닝 기법에 의해 각 질환에 특이적인 혈청 프로테옴의 표본(프로테옴 패턴)을 생성하고, 분석 대상 프로테옴과 각 정상 및 질환의 프로테옴 표본을 비교하여 암 질환 여부 및 그 질환을 특정할 수 있는 방법에 관한 것이다.

<23> 또한, 본 발명은 상기의 암 질환의 진단 방법을 도입한 시스템에 관한 것으로

로, 구체적으로는 이미지 마이닝 기법에 의해 혈청 프로테옴의 표본을 생성하고 프로테옴 표본과 분석 대상 프로테옴을 비교하여 암 질환 여부를 확인할 수 있는 암 질환 분석 시스템 및 방법에 관한 것이다.

<24> 또한, 본 발명은 상기의 암 진단 분석 시스템 또는 방법에 적용할 수 있는 특정의 1이상의 혈청 단백질을 포함하는 암 특이적 진단 마커인 암 진단 프로테옴 패턴을 제공한다.

<25> 근래에는 생물정보학 및 DNA 염기서열 분석기술의 발달로 인간, 동식물, 미생물에 대한 방대한 게놈 데이터가 밝혀졌으며, 이에 따라 의약품 개발, 질병 진단, 작물 육종 등 여러 가지 산업에 새로운 가능성을 열어줄 것으로 기대되고 있다. 생물정보학은 바이오 기술(BT)과 정보 기술(IT) 분야의 학제간 연구를 통해 방대한 양의 데이터를 빠르고 효율적으로 가공하고 처리하는 기술의 하나로, 생명체가 지니고 있는 광대한 양의 정보를 수집, 저장, 분석하고 그 결과를 제약, 식품, 농업, 환경 등의 분야에 이용하여 부가가치를 창출하는 기술이라 할 수 있다.

<26> 인간 게놈 데이터의 완성과 생물정보학의 발전에 따라, 유전자가 인간의 질병 또는 형질에 결정적인 영향을 미치고 있다는 사실이 확인되었다. 즉, 인간은 거의 같은 DNA 서열을 가지고 있지만 외모, 키, 성격 등에 있어서 차이를 보이고 있으며, 외부의 환경적인 요인이 없다면 각각의 인간 게놈이 사람의 특징을 결정한다는 것이다.

<27> 이와 같이, 인간 게놈과 이를 이용한 임상 데이터를 이용하면 암과 같은 난치병을 치료할 수 있을 것으로 전망되며, 암과 같은 질병의 경우 조기에 발견할 수 있다면 더욱 높은 치료율을 기대할 수 있을 것이다. 질병을 조기에 발견하기 위한 시료로서 뇨, 눈물, 타액 등이 사용되고 있으나, 최근에는 혈청 프로테옴이 많이 사용되고 있다.

- <28> 난치성 암질환은 유전적 요인과 환경적 요인이 복합적으로 작용하는 질환으로, 이의 진단과 예후 판정을 위해서는 질환의 발생, 진행 및 악성화와 함께 수반되어 나타나는 총체적 프로테옴 분석이 필요하다. 1~2 종류의 세포나 조직이 아니라, 여러 장기의 기능과 프로세스에 의해 발현되는 암 질환에서 나타나는 프로테옴 변화를 가장 민감하게 반영할 수 있는 생체 시료는 혈청과 같은 체액이라 할 수 있다. 혈청은 확보가 용이하고 일상적인 임상검사에서 분석되는 시료이므로 폐암과 같이 조기진단이나 예후 판정이 어려운 질환에서 우선적으로 고려되어야 하는 최적의 시료라 할 수 있다.
- <29> 암 환자와 정상인의 혈청 프로테옴 성분을 분석하면, 그 성분이 상이할 것으로 예측할 수 있다. 그러나, 개개인의 혈청 프로테옴을 구성하는 어떠한 성분이 어떻게, 얼마나 다른지에 대해서는 알려져 있지 않은 현실이다. 이미 인간 게놈에 의한 유전자 지도가 완성되어 있다고는 하나 유전자에 담긴 정보에 의하여 생성되는 모든 단백질과 유전자와의 상관관계에 대한 정보는 아직까지 정확하게 알려져 있지 않다.
- <30> 특히, 암과 같은 질병에 있어서 특정 유전자에 특이한 변이가 발생하며, 유전자 변이에 의한 혈청 프로테옴 성분이 변하게 될 것으로 생각할 수 있다. 이러한 혈청 프로테옴 성분의 변화를 감지함으로써 암과 같은 질병을 조기에 발견할 수 있는 방법에 관한 몇몇 연구들이 있어 왔다. 예를 들면, RARISH, Christopher, Richard 등의 국제출원 PCT/AU01/00877(2001.7.19)에서는 질량 분석법을 통하여 정상적인 사람 또는 동물로부터의 혈청 단백질과 암 환자로부터의 혈청 단백질을 비교하여 변화된 단백질을 찾아내어 이를 암의 마커(marker)로 사용하는 것에 관하여 개시하고 있다. 구체적으로, 암을 가진 사람 또는 동물로부터 혈액 프랙션을 질량 분석법으로 분리하고, 건강한 사람 또는 동물로부터의 혈액 프랙션을 질량 분석법으로 분리하여 양자의 분자 프로파일(profile)

을 비교하여, 정상적인 사람 또는 동물로부터의 분자 프로파일과는 상이한 암을 가진 사람 또는 동물로부터의 분자를 확인함으로써 증가되거나 또는 감소된 수준의 분자가 암 마커로 확인하는 방법에 대하여 개시하고 있다.

<31> 또한, YIP, Tai-Tung 등의 국제출원 PCT/US01/28133(2001.9.7)에서는 유방암을 진단하는 질량 분석법(SELDI)을 사용하여 찾아낸 새로운 단백질 마커를 공개하고 있다. 이 발명에서는 정상적인 환자와 암 환자의 구별은 단백질 마커의 양, 존재 유무, 발견 빈도 등을 고려하여 수행될 수 있음을 개시하고 있다.

<32> 최근의 Emanuel F. Petricoin III 등의 연구(Lancet, 359;572-577, 2002)에서는 질량 분석법(SELDI-TOF)을 이용하여 정상적인 상태와 구별되는 난소암 환자의 프로테옴 패턴을 찾아내고, 이를 이용하여 높은 정도의 감도 및 특이성으로 난소암 여부를 진단한 바 있다.

<33> 하지만, 상기와 같은 연구들은 모두 질량 분석법(SELDI-TOF 또는 MALDI-TOF)을 이용하여 1차원의 분석패턴으로 정상인의 혈청 단백질과 구별되는 암 환자의 혈청 단백질을 찾아내어 이를 암 질환 마커로 활용한 것인데, 이는 정상인의 혈청 단백질과 구별되는 암 환자의 혈청 단백질을 구별하기 위해서는 질량이라는 하나의 요소(factor)만을 판단의 기준으로 하고 있으며, 수많은 혈청 단백질의 변화 양상과 그 변화 정도를 고려하여 선별하여야 하는 바, 암 진단의 정확성이 떨어질 뿐 아니라 경제적이지 못한 면이 있었다.

<34> 이에 본 발명자들은 상기의 연구의 문제점을 극복하고, 질환의 진단의 정확성을 높이면서도 경제적이며, 의사뿐 아니라 환자에게도 빠르고 간편한 방법으로 누구나 이용할 수 있는 질환 진단 방법을 연구한 결과, 혈청 시료를 2차원으로 분리시킴으로써 변화된

단백질을 더 용이하게 구분할 수 있도록 혈청 단백질의 2D-Gel 분리 상태를 이미지로 변환하여, 이미지 마이닝 기법에 의해 각 질환에 특이적인 혈청 단백질의 표본(단백질 패턴)을 생성하고, 이를 이용하여 암 질환을 진단하고 분석할 수 있는 간편하고 경제적인 방법 및 시스템을 제공하게 되었다.

【발명이 이루고자 하는 기술적 과제】

<35> 본 발명은 상술한 문제점을 해결하기 위하여 안출된 것으로서, 정상적인 개체 및 질환을 가진 개체의 다수의 혈청 단백질을 채취하고 2D-Gel를 이용하여 2차원으로 이미지화 한 후, 이를 이미지 마이닝 기법을 적용하여 혈청 단백질 표본으로 데이터베이스화하고, 분석 대상 혈청 단백질을 입력받아 이미지화한 후 데이터베이스에 저장된 다수의 혈청 단백질 표본과 비교하여 암 질환 여부를 용이하게 판단할 수 있는 단백질 이미지 마이닝을 이용한 암 질환 분석 방법 및 이를 도입한 시스템을 제공하는데 그 기술적 과제가 있다.

<36> 본 발명의 다른 기술적 과제는 2차원으로 이미지화한 혈청 단백질에 이미지 마이닝 기법을 적용함으로써 질환을 가진 개체로부터의 정상 개체와 구별되는 혈청 단백질의 특징적인 패턴을 찾아내는데 용이한 방법을 제공하고자 한다.

<37> 또한, 본 발명의 다른 기술적 과제는 혈청 단백질을 유전자 알고리즘 및 서포트 벡터 머신을 통한 이미지 마이닝 기법으로 분석함으로써 보다 정확한 분석 결과를 얻을 수 있고, 퍼지 룰(fuzzy rule) 기반 분류 과정에 의해 질환의 진행 상태 및 예후를 추적할 수 있도록 한 단백질 이미지 마이닝을 이용한 암 질환 분석 시스템 및 방법을 제공하는 데 있다.

<38> 또한, 본 발명의 기술적 과제는 상기 암 진단 방법 또는 암 진단 시스템에 적용하여 효과적으로 암 진단을 가능하게 하는 암 특이적 진단 마커 즉, 암 진단 프로테옴 패턴을 제공하는데 있다.

【발명의 구성 및 작용】

<39> 상술한 기술적 과제를 달성하기 위한 본 발명의 암 질환 분석 방법은 정상적인 개체 및 암 질환 개체의 혈청 프로테옴을 입력받아 2차원 이미지 처리를 수행하고, 각각의 특징 데이터를 추출한 후, 상기 각각의 특징 데이터로부터 양자를 구별할 수 있는 최적의 특징을 산출하여 질환에 특이적인 프로테옴 패턴을 가진 프로테옴 표본을 생성하고 데이터베이스화하는 과정; 암 질환의 발병 여부를 확인하고자 하는 분석 대상 혈청 프로테옴을 입력받아 상기 분석 대상 혈청 프로테옴을 2차원 이미지화하고 이로부터 특징 데이터를 추출하고, 상기 분석 대상 혈청 프로테옴의 패턴 구조를 상기 질환 특이적 프로테옴 패턴을 가진 프로테옴 표본과 비교하여, 비교 결과에 따라 분석 대상 혈청 프로테옴이 정상인지 암 질환 상태인지 또는 어떤 질환인지를 확인하고 분석하는 과정을 포함한다.

<40> 또한, 본 발명의 암 질환 분석 시스템은 혈청 프로테옴을 입력받기 위한 입력 수단; 다수의 정상 개체 및 질환 개체로부터 채취한 혈청 프로테옴을 입력받아 각각 2차원 이미지로 변환한 후 특징을 추출하고, 상기 각각의 특징 데이터로부터 양자를 구별할 수 있는 최적의 특징을 산출하여 질환에 특이적인 프로테옴 패턴을 가진 프로테옴 표본을 생성하는 한편, 분석 대상 혈청 프로테옴을 2차원 이미지로 변환하고 특징을 추출하기 위한 프로테옴 표본 생성 수단; 상기 프로테옴 표본 생성 수단에 의해 추출된 분석 대상 혈청 프로테옴의 패턴을 상기 프로테옴 표본의 패턴과 맵핑시켜 유사 여부를 판단

하기 위한 프로테옴 비교 수단; 상기 프로테옴 비교 수단의 비교 결과 분석 대상 혈청 프로테옴의 패턴이 정상인의 혈청 프로테옴 패턴과 유사한 경우에는 분석 대상 혈청 프로테옴을 정상으로 판단하고 그렇지 않을 경우에는 암 질환이 발병한 것으로 판단하기 위한 질환 분석 수단; 및 상기 질환 분석 수단의 분석 결과를 출력하기 위한 출력 수단을 구비한다.

<41> 용어의 정의

<42> 다른 정의가 없다면 이 명세서에서 사용된 기술 용어 및 과학 용어는 이 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 통상적으로 이해하고 있는 의미를 가진다.

<43> 본 발명에서 사용된 "마커"는 정상 개체의 혈청 샘플과 비교하여 어떤 질환을 가진 개체의 혈청 샘플에 구별되게 존재하는 폴리펩타이드를 말한다. 이 마커 또는 마커들은 단일의 폴리펩타이드 또는 폴리펩타이드의 조합으로 이루어질 수 있다. 또한, "구별되게 존재하는"의 의미는 정상인의 혈청 샘플과 비교하여 어떤 질환을 가진 자의 혈청 샘플에서 특정 폴리펩타이드가 증가 또는 감소된 수준으로 존재하거나, 또는 정상인의 혈청 샘플에서 새로 발견 또는 소멸되는 것을 의미한다.

<44> 본 발명에서 사용된 "프로테옴 패턴"은 정상 개체의 혈청 샘플과 비교하여 어떤 질환을 가진 개체의 혈청 샘플에서 구별되게 존재하는 폴리펩타이드의 특징적인 그룹 또는 그룹핑되는 형태를 의미한다. 즉, 특정 질환의 경우에 특이적으로 변화 양상을 나타내는 혈청 단백질의 그룹 또는 이들의 2차원에서의 위치 형태를 등을 포함한다. 또한, 본 발명에서 사용된 "질환 특이적 프로테옴 패턴"은 특정 질환에 따라 특이적으로 나타나

는 혈청 단백질의 그룹 또는 이들의 그룹핑 형태를 말한다. 이는 본 발명의 방법 및 시스템의 질환 판별의 마커로 사용된다.

<45> 본 발명에서 사용된 "특징 데이터"는 정상 개체 및 질환 개체로부터의 혈청 프로테옴의 비교를 통해 질환 상태를 구별해 낼 수 있는 혈청 프로테옴들에 관한 데이터를 말한다. 구체적으로는 질환 상태의 혈청 2차원 이미지에 특이적으로 존재하는 혈청 프로테옴에 해당하는 스팟들의 데이터를 포함한다. 예를 들면, 이는 스팟들의 그룹(조합), 각 스팟들의 질량 및/또는 등전점 등을 포함할 수 있다. 또한, 본 발명에서 사용된 "최적의 특징 데이터"는 상기 특징 데이터 중에서 질환을 특이적으로 판별해 낼 수 있는 최적의 데이터를 말한다. 구체적으로, 질환에 특이적인 스팟들의 조합 중에서 최적의 조합을 포함한다.

<46> 본 발명에서 사용된 "데이터 마이닝"은 수많은 데이터 가운데 숨겨져 있는 유용한 상관 관계를 발견하는 것으로, 데이터베이스로부터 과거에는 알지 못했지만 데이터 속에서 유도된 새로운 데이터 모델을 발견하여 미래에 실행 가능한 정보를 추출해 내고 의사 결정에 이용하는 과정을 말한다. 즉, 데이터에 숨겨진 패턴과 관계를 찾아내어 정보를 발견해 내는 것을 의미한다.

<47> 본 발명에서 사용된 "유전자 알고리즘(GA)"은 자연계에 있어서 생물의 유전과 진화의 메커니즘을 공학적으로 모델화하는 것에 의해 생물이 갖는 환경에서의 적응능력을 취급하는 것으로, 풀고자 하는 문제에 대한 가능한 해들을 정해진 형태의 자료구조로 표현한 다음 이들을 점차적으로 변형함으로써 점점 더 좋은 해들을 생성해 내는 방법이다. 보다 구체적으로, 유전자 알고리즘은 어떤 범위 내에서 정의되어 있는 변수 x 에 대한 함수 $f(x)$ 의 최대치 또는 최소치를 이끌어 내는 x 값을 고속으로 구하기 위한 최적화 탐색

알고리즘의 일종이라고 할 수 있다. 유전자 알고리즘은 유전자 요소를 기호열로 변환하는 코딩 작업을 수행하는 유전자형 결정 단계, 유전자형 결정 단계에서 결정된 유전자형으로부터 요소가 다른 다양한 개체를 발생시키기 위한 초기 유전자 집단 결정 단계, 각 개체의 적응도를 미리 결정한 방법에 의해 연산하기 위한 개체의 적응도 평가 단계, 적응도 평가 단계에서 결정한 적응도에 기초하여 개체의 생존 분포를 결정하기 위한 선택 단계, 2개의 염색체 사이에서 유전자를 바꾸어 넣어 새로운 개체를 발생시키는 교배처리 단계, 유전자의 어떤 부분의 값을 강제적으로 바꾸고 유전자 집단으로서의 다양성을 극대화하여 보다 좋은 해를 가지는 개체가 발생되도록 하는 돌연변이 처리 단계, 각 개체의 적응도 평가단계로 복귀하는 단계로 이루어진다. 유전자 알고리즘을 이용하게 되면 복수의 개체 사이에서 선택, 교배 등의 유전적 조작에 의해 상호 협력적으로 해의 탐색을 수행하기 때문에 단순한 병렬적 해의 탐색과 비교하여 보다 좋은 해를 발견하기 쉽고, 알고리즘이 단순하여 실시가 용이한 장점이 있다.

<48> 본 발명에서 사용된 "서포트 벡터 머신(SVM)"은 패턴 인식에 유용하게 사용되는 학습용 머신으로서, 결정되는 표면이 서포트 벡터 및 그에 대응하는 가중치의 집합으로 이루어지는 변수에 의해 결정되며, 다수의 변수를 각각 따로 취급하지 않고 동시에 처리하는 방법을 제시한 것을 의미하고, 이는 벡터를 분류하는 유용한 도구로 사용될 수 있다. 서포트 벡터 머신에 의하면 입력 공간의 비선형적인 높은 차수를 특징면(feature space)에서 선형적으로 투영하여 해석할 수 있도록 하며, 각 특징 사이의 최적의 경계(최적 분리면)을 제시한다. 서포트 벡터 머신은 크게 트레이닝 과정과 테스트 과정의 두 부분으로 이루어진다. 트레이닝 과정에서는 서포트 벡터가 생성되며 테스트 과정에서는 특정 규칙에 의한 판단이 수행된다.

<49> 이하에서는 본 발명의 질환 분석 방법 및 이를 이용한 시스템에 대하여 첨부된 도면을 참조하여 구체적으로 설명한다. 이하에서 설명하는 질환 분석 시스템 및 방법은 다양한 질환에 적용될 수 있지만 본 발명에서는 특히 암 질환에 적용하는 경우를 예로 들어 설명하기로 한다.

<50> 또한, 표본 생성 및 질환 분석을 위해 사용되는 시료는 혈청, 뇨, 눈물 타액 등 정상적인 상태와 구별될 수 있는 질환 특이적 폴리펩타이드를 확인할 수 있는 생체 시료를 포함한다. 또한, 구체적으로 혈청 프로테옴은 유전자를 갖는 모든 개체에 적용 가능하지만 본 발명에서는 사람인 경우를 예로 들어 설명한다.

<51> 본 발명에 있어서, 암 질환은 '통제할 수 없는 세포 성장'으로 인한 병적 상태를 의미하는 것으로, 구체적인 예를 들자면 유방암, 난소암, 위암, 간암, 자궁암, 폐암, 대장암, 췌장암, 전립선암 등을 포함한다.

<52> 이미지 마이닝 기법을 이용한 질환 분석 시스템

<53> 도 1은 본 발명에 의한 암 질환 분석 시스템의 구성도이다.

<54> 도시된 것과 같이, 본 발명에 의한 암 질환 분석 시스템(10)은 프로테옴 표본 생성 수단(102), 프로테옴 비교 수단(104), 질환 분석 수단(106), 입출력 인터페이스(108), 제어부(110), 입력 수단(112), 출력 수단(114) 및 데이터베이스(116)를 포함한다.

<55> 프로테옴 표본 생성 수단(102)은 적어도 N명(예를 들어, 20명)의 정상 개체 및 N명(예를 들어, 20명)의 질환을 가진 개체로부터 채취한 혈청 프로테옴을 입력 수단(112)을

통해 입력받아 2차원 이미지로 변환한 후(도 5 참조) 특징 즉, 질환에 특이적인 스팟들을 추출하고, 추출된 특징 데이터 중에서 최적의 특징 데이터를 식별해 내는 한편, 2차원 이미지를 구성하는 스팟으로부터 데이터 간의 상호 관계성을 추출하고 규칙화하여 데이터베이스(116)에 저장한다. 이를 위하여 프로테옴 표본 생성 수단(102)은 유전 알고리즘, SVM, 퍼지 룰 기반 분류 등의 기법을 사용할 수 있으며, 이러한 기법들에 대한 구체적인 설명은 후술할 것이다. 이와 같이 하여, 정상인의 혈청 프로테옴 표본과 구별되는 암 질환 개체의 혈청 프로테옴의 특징(농도, 크기 등)을 알아낼 수 있게 되며, 특히 퍼지 룰 기반 분류 기법을 사용하는 경우에는 암 질환 여부뿐만 아니라 질병의 진행 정도, 예후 추적 등이 가능하게 된다.

<56> 또한, 프로테옴 표본 생성 수단(102)은 표본의 대상이 되는 정상인 및 질환개체의 혈청 프로테옴 뿐만 아니라 분석을 원하는 대상의 혈청 프로테옴을 2차원 이미지로 변환하고 특징 데이터를 추출하여, 이후의 질환 여부를 분석하는 과정에서 사용될 수 있도록 한다.

<57> 프로테옴 비교 수단(104)은 프로테옴 표본 생성 수단(102)에 의해 분석 대상 혈청 프로테옴의 특징 데이터가 추출되면 분석 대상 혈청 프로테옴의 패턴과 데이터베이스(116)에 저장된 프로테옴 표본의 패턴을 맵핑시켜 유사한지의 여부를 판단한다.

<58> 질환 분석 수단(106)은 분석 대상 혈청 프로테옴의 패턴이 정상인의 혈청 프로테옴과 유사한 경우에는 분석 대상 혈청 프로테옴을 정상으로 판단하고 그렇지 않을 경우에는 암 질환이 발병한 것으로 판단하며, 판단 결과를 출력 수단(114)으로 출력한다. 이때, 암 질환 개체로 판명된 경우에는 그 진행 정도, 예후 등을 예측하여 출력하며, 암

질환개체가 아닌 경우에도 향후 암 질환 발병 가능성이 있는지, 가능성이 있다면 얼마나 되는지에 대해서도 예측하여 출력하는 것도 가능하다. 이를 위하여 프로테옴 표본 생성 수단(102)에서 퍼지 룰 기반 분류 기법을 이용하여 표본 데이터를 생성하여야 한다.

<59> 입출력 인터페이스(108)는 암 질환 분석 시스템(10)과 외부 장치와의 접속 및 연동을 위한 것이며, 제어부(110)는 이상에서 설명한 각 기능 수단의 전체적인 동작을 제어한다.

<60> 도시하지 않았지만, 본 발명의 암 질환 분석 시스템(10)은 암호화 수단을 더 구비하여, 혈청 프로테옴을 제공한 정상인 및 암 질환개체의 신상 정보, 분석 대상 혈청 프로테옴 제공자의 신상 정보가 암호화된 상태로 데이터베이스(116)에 저장될 수 있도록 한다.

<61> 도 2는 도 1에 도시한 프로테옴 표본 생성 수단(102)의 상세 구성도이다.

<62> 프로테옴 표본 생성 수단(102)은 혈청 프로테옴의 2차원 이미지로부터 의미 있는 특징 데이터를 획득하기 위한 전처리 수단(210), 전처리 수단(210)에서 획득한 특징 데이터로부터 분석 대상 혈청 프로테옴의 정상 여부를 확인하기 위한 진화적 분류 수단(220) 및 전처리 수단(210)에서 획득한 특징 데이터로부터 경험적 지식, 통계적 방식 등을 통하여 분석 대상 혈청 프로테옴의 보다 세밀한 상태를 판단하기 위한 퍼지 룰 기반 분류 수단(230)을 구비한다.

<63> 전처리 수단(210)은 또한 이미지 처리 수단(212) 및 특징 추출 수단(214)을 포함하는데, 이미지 처리 수단(212)은 입력된 2차원 이미지로부터 잡음 제거(noise filtering), 이미지 증폭(image enhancement), 직교 투영(ortho-projection), 예지

(edge) 검출, 최적 스레쉬홀딩(optimal thresholding) 등 기본적인 영상 처리 작업을 수행하고, 특징 추출 수단(214)은 영상 처리된 2차원 이미지로부터 기본적인 특징 즉, 질환에 특이적인 스팟들을 추출한다. 특징 추출 수단(214)에서 추출한 각 특징들은 식별화 또는 라벨링됨으로써 스팟별 특징 데이터가 생성된다.

<64> 진화적 분류 수단(220)은 전처리 수단(210)에서 획득한 데이터를 사용하여 정상 또는 질환을 갖는 개체의 혈청 프로테옴 패턴을 분석하는 수단으로서, GA(Genetic Algorithm) 처리 수단(222) 및 SVM 적용 수단(224)을 포함하여 질환에 특이적인 스팟의 조합 중 최적의 조합을 찾아내는 수단이다. GA 처리 수단(222)은 전처리 수단(210)에서 추출된 특징 데이터(질환에 특이적인 스팟)들 중에서 분류에 결정적인 역할을 담당할 최적의 특징들을 식별하고, SVM 적용 수단(224)은 판단 함수 및 분류 오차율을 이용하여 GA 처리 수단(222)에서 식별해 낸 최적의 특징 데이터들의 적합도를 판단한다. 이에 의하여 다음 세대의 스팟에 대한 대안이 생성되고, 이러한 진화방법을 통하여 최적의 특징 데이터 및 그에 따른 판단 함수가 생성될 수 있다. 여기에서, SVM 적용 수단(224)에서 사용하는 판단 함수는 이미 정해진 함수이다.

<65> GA는 최근에 각광받는 최적화 기법으로, 본 발명에서는 GA를 이용하여 5개 내지 20개 정보의 특징들을 효과적으로 추출할 수 있다. 진화적 분류 수단(220)의 진화적 분류 기법에 의해 암 등의 질병 여부를 용이하고도 효율적으로 분석할 수 있는 복수개의 특징들이 추출된다. 향후 분석 대상개체로부터 채취한 테스트 샘플로부터 추출되는 특징들은 상기 복수개의 특징들과 비교됨으로써 피검자가 암 등의 질병에 감염되었는지의 여부를 판별할 수 있다.

<66> 한편, 퍼지 룰 기반 분류 수단(230)은 진화적 분류 과정에서 간과될 수 있는 가공된 정보들(가족의 병력, 현재 자신의 병력 등), 예를 들어 특정 스팟 간의 상호 관계성 등을 통계적, 경험적 방법을 통해 추출해 냄으로써, 보다 상세한 분류 및 인식 정확도를 향상시키기 위한 것으로, 데이터 맵핑 수단(232) 및 룰 기반 분류 수단(243)을 포함한다. 데이터 맵핑 수단(232)은 혈청 프로테옴의 2차원 이미지로부터 스팟 간의 상호 관계성을 산출하고, 산출된 특성들을 통계적 기법에 의해 분류하며, 퍼지 기법을 이용하여 통계적 불확실성을 정량화한다. 룰 기반 분류 수단(234)은 데이터 맵핑 수단(232)의 결과를 정리하고 규칙화하여 최종적으로 룰 베이스화한다. 퍼지 룰 기반 분류 수단(230)은 본 발명에서 있어서 반드시 필요한 수단은 아니지만, 이를 적용할 경우 암 질환의 단순한 판단만이 아니라, 전문가 시스템에 의한 통계적, 경험적 방법에 의하여 질병의 진행 정도, 예후 추적 등을 가능하게 한다.

<67> 다음으로, 본 발명의 질환 분석 방법에 대하여 설명하기로 한다. 본 발명에 의한 암 질환 분석 방법은 질환에 특이적인 프로테옴 패턴을 가진 프로테옴 표본을 생성하고 데이터베이스화하는 과정(트레이닝 단계); 및 분석을 원하는 대상 혈청 프로테옴의 특징 데이터를 추출하여 상기 질환 특이적 프로테옴 표본과의 비교를 통한 대상 혈청 프로테옴의 질환 여부를 분석하는 과정(테스팅 단계)을 포함한다. 이하 구체적으로 설명한다. 본 발명에 따른 암 질환 분석 방법은 이러한 방법을 수행하기 위한 프로그램이 저장되어 있는 메모리와 상기 메모리에 결합된 프로세서에 의해 실행될 수 있다. 또한, 디지털 처리 장치에 의하여 실행될 수 있는 명령어들의 프로그램이 유형적으로 구현되어 있으며 디지털 처리 장치에 의하여 판독될 수 있는 기록 매체에 본 발명에 따른 진화적 분류 과

정 및 퍼지 를 기반 분류 과정의 적용에 의한 질환 확인 방법에 대한 프로그램이 저장될 수 있다. 아울러, 본 발명에 따른 암 질환 분석 방법, 이 방법을 실현하기 위한 각각의 모듈들은 소프트웨어, FPGA, ASIC 등으로 구현될 수 있음은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게는 자명할 것이다.

<68> 질환 특이적 프로테옴 패턴을 가진 프로테옴 표본의 생성(트레이닝 단계; S1)

<69> 도 3은 본 발명에 의한 암 질환 분석 방법을 설명하기 위한 흐름도이고, 도 4는 도 3에 도시한 프로테옴 표본 생성 방법을 설명하기 위한 흐름도이다.

<70> 본 발명의 질환 특이적 프로테옴 표본의 생성은 정상적인 개체로부터의 혈청 프로테옴과 특정 질환을 가진 개체로부터의 혈청 프로테옴을 비교하여 특정 질환에 특이적인 프로테옴 패턴을 찾음으로써 수행된다.

<71> 본 발명에서는 혈청 프로테옴의 표본을 결정하기 위하여 이미지 데이터 마이닝 기법을 이용하며, 보다 구체적으로는 서포트 벡터 머신(Support Vector Machine)을 적용한다. 데이터 마이닝은 수많은 데이터 가운데 숨겨져 있는 유용한 상관관계를 발견하는 것으로, 데이터베이스로부터 과거에는 알지 못했지만 데이터 속에서 유도된 새로운 데이터 모델을 발견하여 미래에 실행 가능한 정보를 추출해 내고 의사 결정에 이용하는 과정을 말한다. 즉, 데이터에 숨겨진 패턴과 관계를 찾아내어 정보를 발견해 내는 것이다. 데이터 마이닝은 이미지 분석에 적용될 수 있으며, 이미지 분석은 디지털화된 사진으로부터 패턴을 추출하는 기법으로 문자인식, 의료진단, 방위산업 등 다양한 분야에서 이용되고 있다.

<72> 질환 특이적 프로테옴 패턴을 찾는 방법은 정상 및 질환 개체로부터의 혈청 프로테옴을 입력 수단(112)을 통해 입력받아(S101) 2D-Gel를 사용하여 분리하고, 프로테옴 표본 생성 수단(102)에서 이를 2차원 이미지 정보로 전환하여 질환(특히, 암)에 특이적인 특징을 추출한 후, 추출된 특징 데이터 중에서 최적의 특징 데이터를 찾는 최적화하여 표본을 생성하고 데이터베이스(116)에 저장하는 과정(S102)을 포함한다. 프로테옴 표본 생성 과정을 도 4내지 도 6을 참조하여 보다 상세히 설명하면 다음과 같다.

<73> 현재까지 프로테옴의 분석은 다양한 방법을 통해 수행되고 있지만, 본 발명에서는 바람직하게 2D-Gel 방법을 사용한다. 본 발명에서의 2D-Gel 분석방법은, 이 기술 분야에서 일반적으로 사용되는 방법에 따라 수행되는데, 단백질이 가지는 전하량에 따라 단백질을 1차 분리한 후(IEF, isoelectric focusing), 분자량에 따라 (SDS-PAGE) 2차 분리하는 원리를 이용한 것이다. 본 발명에서는 정상 개체로부터의 혈청 단백질 및 질환을 가진 개체로부터의 혈청 단백질을 각각 2D-Gel를 사용하여 분리한다.

<74> 도 5는 혈청 프로테옴 이미지의 일 예를 나타내는 사진이며, 상기와 같이 2D-Gel로 분리된 혈청 단백질들을 도 5에 도시한 것과 같이 디지털화된 사진 등으로 변환시킴으로써 이미지 분석이 가능하도록 가공한다. 이후, 디지털 정보 형태로 전환된 상기 혈청 프로테옴의 이미지 정보로부터 질환에 특이적인 특징을 추출하여 저장한다. 즉, 다수의 정상적인 개체 및 질환을 가진 개체로부터의 혈청 프로테옴의 2차원 이미지 정보들에서 특징적인 공통되는 특징을 추출하고, 이 특징들이 갖는 각각의 데이터(좌표, 분자량, 등전점 등을 포함)를 저장하여 데이터베이스화한다. 구체적으로 예를 들면, 질환을 가지는 개체로부터의 혈청 프로테옴 2차원 이미지에서 정상적인 개체로부터의 혈청 프로테옴의 2차원 이미지와 구별되는 스팟의 농도(intensity), 크기(size) 등을 가진 스팟들 또는

정상적인 경우에는 존재하지 않는 새로운 스팟(new spot) 등에 대한 정보(좌표, 분자량, 등전점 등을 포함)를 저장하여 데이터베이스화한다. 또한, 다수의 질환을 가진 개체로부터의 혈청 프로테옴의 2차원 이미지 정보들에서 공통적인 특징으로 추출하고 이 특징들이 갖는 각각의 데이터(좌표, 분자량, 등전점 등을 포함)를 데이터베이스화한다.

<75> 여기서, 질환에 특이적인 특징이란 도 5에 도시한 것과 같이 전하량과 분자량에 따라 분리한 혈청 프로테옴 이미지의 스팟들 중 질환에 특이적인 농도 및 크기를 갖는 특정 스팟을 의미한다.

<76> 구체적으로 예를 들어 설명하면, 혈청 프로테옴의 분석을 통하여 수많은 단백질의 분자량 및 산도를 측정하여 각각에 대하여 번호를 부여하고, 이 중 특정 암을 효과적으로 진단할 수 있는 암 진단 마커인 암 특이적 프로테옴 패턴으로 추출된다. 유방암 환자의 경우의 예로 하기 표1에서는 환자의 수많은 혈청 프로테옴으로부터 유방암에 특이적인 특징을 나타내는 총 67개의 특정 스팟들을 선별하고, 이들 스팟들은 분자량 및 등전점(pI)로서 특정된다.

<77>

【표 1】

스팟 번호	질량(kDa)	등전점(pI)	4319	42.9	5.59
106	20.2	4.60	4321	47.1	5.63
204	27.4	4.57	4401	48.2	5.40
406	52.8	4.58	4601	74.0	5.43
1109	15.7	4.87	5108	26.2	5.78
1303	41.0	4.82	5122	21.0	5.71
1311	39.9	4.89	5303	42.9	5.71
1318	38.4	4.99	5304	39.3	5.71
2105	20.0	5.04	5305	40.9	5.71
2114	25.8	5.22	5313	40.7	5.85
2115	16.2	5.19	5314	41.8	5.88
2204	32.9	5.07	5405	50.6	5.73
2301	41.0	5.02	5406	56.2	5.75
2309	46.4	5.13	5408	52.2	5.75
2316	42.1	5.21	5409	47.7	5.76
2318	44.8	5.21	5417	54.3	5.68
2705	138.3	5.03	6113	20.5	6.05
3104	17.8	5.23	6115	21.0	6.08
3106	24.9	5.23	6204	35.5	5.92
3110	22.8	5.29	6231	33.9	6.08
3113	25.8	5.32	6302	42.5	5.89
3119	21.0	5.39	6305	47.1	5.97
3306	40.7	5.23	6306	39.4	5.98
3402	51.2	5.22	6307	45.6	5.99
3421	57.3	5.23	6308	37.7	5.99
3514	62.7	5.38	6309	41.6	6.03
4110	25.7	5.49	6310	46.3	6.04
4123	17.7	5.50	6311	44.6	6.04
4215	28.6	5.51	6312	45.6	6.08
4302	39.4	5.42	6403	52.0	5.89
4311	40.7	5.50	6418	53.1	6.05
4313	37.3	5.53	7307	38.0	6.12
4318	44.9	5.59	7312	42.7	6.14
			7314	44.1	6.18

<78> 상기의 각각 추출하여 저장된 특징 데이터 중에서 최적의 특징 데이터를 식별해 내는 한편, 2차원 이미지를 구성하는 스팟(spot)으로부터 데이터간의 상호 관계성을 추출하고 규칙화하여 데이터베이스화한다. 이를 위하여, 유전 알고리즘, 서포트 벡터 머신(Support Vector Machine; 이하, SVM 이라 함), 퍼지 를 기반 분류 등의 기법을 사용할 수 있다. 이와 같이 하여 검출된 최적의 특징 데이터는 정상적인 개체와 구별되는 특징의 질환을 가진 개체의 프로테옴 패턴, 즉 질환 특이적 프로테옴 패턴이 되는 것이다.

<79> 구체적으로 상기 표1의 스팟들의 조합은 유방암의 특이적 프로테옴 패턴의 한 형태가 될 수 있는데, 본 발명에 있어서 유방암 진단의 경우에는 상기 표1의 스팟들로부터 1 이상이 선택된 스팟들의 조합이 유방암의 특이적 패턴으로 사용될 수 있다. 여기서 1 이상이 선택된다는 것은 총 67개의 스팟들의 1이상의 조합들, 즉 $\sum_{i=1}^{67} C_n$ 을 의미한다.

<80> 도 4를 참조하여 보다 구체적으로 설명하면, 본 발명의 프로테옴 표본 생성 과정은 혈청 프로테옴의 2차원 이미지로부터 의미 있는 특징 데이터를 획득하기 위한 전처리 과정(S201), 진화적 분류 과정(S202~S204)을 포함하며, 암 질환의 단순한 판단만이 아닌 통계적, 경험적 방법에 의하여 질병의 진행 정도, 예후 추적 등을 가능하게 하기 위하여 퍼지 를 기반 분류 과정(S205~S207)을 더 포함할 수 있다.

<81> 또한, 전처리 과정(S201)은 이미지 처리 과정 및 특징 추출 과정을 포함하는데, 상기 이미지 처리 과정에서는 이미지 처리 수단(212)에 의해 입력된 2차원 이미지로부터 잡음 제거, 이미지 증폭, 직교 투영, 에지 검출 등 기본적인 영상 처리 작업을 포함하고, 상기 특징 추출 과정에서는 특징 추출 수단(214)에 의해 영상 처리된 2차원 이미지로부터 스팟 형태의 기본적인 특징들을 추출한다. 또한, 상기 특징 추출 과정에서 추출한 각 특징들은 식별화 또는 라벨링됨으로써 스팟별 특징 데이터가 생성된다.

<82> 진화적 분류 과정은 전처리 과정에서 획득한 데이터를 그대로 사용하여 정상 또는 질환을 갖는 개체의 혈청 프로테옴 패턴을 분류하는 단계로서, GA(Genetic Algorithm) 처리 과정(S202), SVM 적용 과정(S205), SVM 적용 과정의 결과에 따라 식별된 최적의 특징 데이터들 및 그에 따른 판단 함수를 추출하는 과정(S204)을 포함한다. 상기 GA 처리

과정(S202)은 GA 처리 수단(222)에 의해 전처리 과정에서 추출된 특징 데이터들 중에서 질환에 특이적인 스팟의 분류에 결정적인 역할을 담당할 최적의 특징을 갖는 스팟을 식별한다. 또한, SVM 과정(S203)에서는 SVM 적용 수단(224)에 의해 판단 함수의 분류 오차를 이용하여 GA 처리 과정에서 구별해 낸 최적의 특징 데이터들의 적합도를 판단한다. 이에 의하여 다음 세대의 스팟에 대한 대안이 생성되고, 이러한 진화방법을 통하여 최적의 특징 데이터 및 그에 따른 판단 함수가 생성될 수 있다(S204). 여기에서 SVM 적용 과정에서 사용하는 판단 함수는 이미 정해진 함수이다.

<83> 도 6은 전처리 과정으로부터 진화적 분류 과정까지를 도식화한 것으로, 질환을 갖는 개체의 혈청 프로테옴 이미지 집합(100) 및 정상 개체의 혈청 프로테옴 이미지 집합(200)으로부터 이미지 전처리 과정을 수행한 후(300), 질환을 갖는 개체 및 정상 개체의 프로테옴 이미지로부터 질환에 특이적인 스팟들을 추출하여, 각 스팟들이 갖는 농도 및 크기에 따라 질환 여부를 결정하여 데이터베이스(400)화한다. 이와 같이 하여 추출된 다수의 특징들은 5 이상, 바람직하게는 5 내지 100개의 특징 스팟을 갖고 있으며, 이 1세대 특징 데이터(질환에 특이적인 스팟)들을 SVM에 적용하여(500) 최적의 특징 데이터 및 그에 따른 판단 함수가 생성되게 된다. 또한, GA에 의해 유전자의 교배, 돌연변이를 발생시킨 결과 즉, 다음 n세대까지에 대한 혈청 프로테옴 이미지에 대해서도 동일한 과정을 수행하여(600, 700) 최종적인 최적 특징 및 판단함수가 도출되게 된다.

<84> 도 7은 서포트 벡터 머신에 의해 최적 분리면이 결정된 상태를 도식화한 것으로, 혈청 프로테옴 이미지로부터 추출된 스팟 1로부터의 특징, 스팟 2로부터의 특징 및 스팟 3으로부터의 특징들의 상호 관계에 의해 정상과 질환을 구별하는 최적의 경계 즉, 최적 분리면이 도출된 예를 나타낸다.

<85> 퍼지 를 기반 분류 과정은 진화적 분류 과정에서 간과될 수 있는 가공된 정보들, 예를 들어 특정 스팟 간의 상호 관계성 등을 통계적, 경험적 방법을 통해 추출해 냄으로써, 보다 상세한 분류 및 인식 정확도를 향상시키기 위한 것으로, 데이터 맵핑 과정(S205), 를 기반 분류 과정(S206) 및 이를 바탕으로 를 베이스를 생성하는 과정(S207)을 포함한다. 데이터 맵핑 과정(S205)은 데이터 맵핑 수단(232)에 의해 혈청 프로테옴의 2차원 이미지로부터 스팟 간의 상호 관계성을 산출하고, 산출된 특성들을 통계적 기법에 의해 분류하며, 퍼지 기법을 이용하여 통계적 불확실성을 정량화한다. 를 기반 분류 과정(S026)에서는 를 기반 분류 수단(234)에 의해 데이터 맵핑의 결과를 정리하고 규칙화하며, 이에 따라 최종적인 를 베이스가 생성되게 된다(S207). 퍼지 를 기반 분류 과정은 본 발명에 있어서 반드시 필요한 수단은 아니지만, 이를 적용할 경우 암 질환의 단순한 판단만이 아니라, 전문가 시스템에 의한 통계적, 경험적 방법에 의하여 질병의 진행 정도, 예후 추적 등을 가능하게 한다.

<86> 본 발명의 프로테옴 표본 생성 과정은 각각 N명(예를 들어, 20명)의 정상인 및 N명(예를 들어, 20명)의 암 질환개체의 혈청 프로테옴의 이미지 정보로부터 특징(질환에 특이적인 스팟)을 추출한 후, 특징 데이터로부터 최적의 특징을 산출하여 프로테옴 표본을 생성하는데, 혈청 프로테옴의 2차원 이미지로부터 경험적 지식, 통계적 방식 등을 통하여 분석 대상 혈청 프로테옴의 보다 세밀한 상태를 판단하는 과정을 더 수행하는 것도 가능하다.

<87> 도8은 질환 특이적 프로테옴 패턴을 가진 프로테옴 표본의 생성(트레이닝 단계;S1) 과정을 유방암의 경우를 실제로 수행하는 경우의 예를 나타낸 것이다. 30명의 정상적인 개체로부터의 혈청 2D 이미지 및 30명의 특정 암 질환 개체로부터의 2D 이미지 분석을

통한 스왑들의 정보를 수집 및 가공하고 SVM/GA를 이용하여 암 질환 특이적 프로테옴 패턴을 탐색한 과정을 보여주고 있다.

<88> 질환 특이적 프로테옴 표본과 분석을 원하는 대상 프로테옴의 비교를 통한 질환 여부 판단(테스팅 단계; S2)

<89> 상기에서와 같이 정상 및 질환을 갖는 개체에 대한 혈청 프로테옴을 분석하여 표본을 생성한 후, 분석을 원하는 대상 혈청 프로테옴 또한 입력 수단(112)을 통해 입력하여(S103) 프로테옴 표본 생성 수단(102)에서 특징 데이터 추출 과정을 수행한다(S104). 보다 상세히 설명하면, 상술한 표본 생성을 위한 이미지 전처리 단계에서와 같이 2D-GeI를 이용하여 분리되고, 이 2차원 이미지를 디지털 정보 형태로 전환된다. 이렇게 입력된 분석 대상 2차원 이미지에 대하여 잡음 제거, 이미지 증폭, 직교 투영, 에지 검출 등 기본적인 영상 처리 작업을 거쳐 특징적인 데이터인 프로테옴 패턴을 추출한다. 이렇게 추출된 프로테옴 패턴은 상기의 질환 특이적 프로테옴 표본들과의 비교를 위해 사용된다.

<90> 본 발명의 질환 특이적 프로테옴 표본과 분석을 원하는 대상 프로테옴의 비교를 통한 질환 여부 판단은 프로테옴 비교(S105)하는 과정 및 암 질환 여부를 확인(S106)하는 과정을 포함한다. 암 질환 분석 결과는 이후 출력 수단(114)을 통해 디스플레이된다(S107).

<91> 프로테옴을 비교(S105)하는 과정은 프로테옴 비교 수단(104)에서 분석 대상 혈청 프로테옴의 패턴 구조를 데이터베이스(116)에 저장된 질환 특이적 프로테옴 표본과 비교

하고, 질환 분석 수단(106)에서 분석 대상 혈청 프로테옴이 정상인지 암 질환 상태인지 확인하고 분석한다. 이 때, 트레이닝 단계에서 퍼지 룰 기반 분류 수단을 이용하여 혈청 프로테옴의 2차원 이미지로부터 경험적 지식, 통계적 방식 등을 통하여 분석 대상 혈청 프로테옴의 보다 세밀한 상태를 판단하여 저장해 둔 경우에는 분석 대상 혈청 프로테옴의 현재 상태뿐만 아니라 예후까지 예측할 수 있게 된다.

<92> 암 진단 여부를 확인하기 위하여 패턴 매칭 과정을 수행하는데, 트레이닝 단계에서 퍼지 룰 기반 분류 수단을 적용한 경우에는 상세 분류 과정을 더 포함할 수 있다. 패턴 매칭 과정은 전처리 과정을 수행한 분석 대상 혈청 프로테옴에 프로테옴 표본 생성시 추출된 특징들 및 판단 함수를 적용하여 정상 또는 질환 여부를 분류하는 과정으로, SVM을 이용하여 수행된다. 또한, 상세 분류 과정은 퍼지 룰 기반 분류 과정에서 의하여 생성된 룰 베이스에 전처리 과정을 수행한 분석 대상 혈청 프로테옴을 투영함으로써 스왑 간의 상호 관계성 등 상세한 분류 결과를 추론하는 과정이다.

<93> SVM은 위의 용어의 정의에서 설명한 것과 같이, 트레이닝 과정 및 테스트 과정으로 이루어지는데, 트레이닝 과정에서는 먼저, 트레이닝 집합에서 데이터 벡터를 입력받는다. 본 발명에서는 각각 N명의 정상인 및 암 질환개체의 혈청 프로테옴의 전처리 과정 결과를 입력받는 과정이 이에 해당된다. 이후, SVM은 입력된 데이터 벡터를 다차원 공간상의 값으로 변환하고, 서포트 벡터 및 가중치에 대한 파라미터를 결정한다. 다음으로, 테스트 과정에서는 테스트 집합에서 데이터 벡터를 입력받고, 이를 데이터 매칭에 의해 다차원 공간 상의 값으로 변환한다. 이후, 각 입력 데이터 벡터의 상태를 나타내는 최적 분리면으로부터 분류 신호를 생성한다. 즉, 테스트 집합에서 입력된 데이터 벡터가 정상인지 비정상인지 판단한다.

<94> 도9에서는 질환 특이적 프로테옴 표본과 분석을 원하는 대상 프로테옴의 비교를 통한 질환 여부 판단(테스팅 단계;S2) 과정의 실패를 나타내었다. 상기 트레이닝 단계(S1)에서 생성된 유방암에 대한 결정 모델을 토대로 검진 대상자 테스트 셋(암 환자 33명, 정상인 35명)을 테스트하였다.

<95> 본 발명의 바람직한 실시예에서, 분석 대상 혈청 프로테옴 및 분석 결과를 데이터 베이스(116)에 저장하여 두면 추후 타 프로테옴 분석에 유용하게 사용할 수 있다.

<96> 다음의 실시예는 본 발명의 시스템 및 방법에 의한 실제적인 암 진단 결과를 나타낸 것이다.

<97> <실시예1>

<98> 30명의 유방암 환자의 혈청 프로테옴 및 30명의 정상인의 혈청 프로테옴의 2차원 이미지로부터 트레이닝 과정을 거친 후, 33명의 암 환자 및 35명의 정상인을 상대로 테스트하였다. 결과는 적중률 94.11%, 민감도 100%, 특이도 88.57%를 나타내었다. 이 경우 26개의 스팟을 유방암의 최적의 특징 데이터로 잡았다. 상기 26개의 스팟은 상기 표1의 67개의 유방암 특이적 스팟들로부터 선택되었다. 구체적인 결과는 도10에 나타내었다. 상기에서 적중률은 유방암의 유무를 올바르게 판단한 정도를, 민감도는 양성율 그 자체를 의미하며, 또한 특이도는 다른 종양과의 감별하는 정도를 의미한다.

<99> 이와 같이, 본 발명이 속하는 기술분야의 당업자는 본 발명이 그 기술적 사상이나 필수적 특징을 변경하지 않고서 다른 구체적인 형태로 실시될 수 있다는 것을 이해할 수 있을 것이다. 그러므로 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적인 것이 아닌 것으로서 이해해야만 한다. 본 발명의 범위는 상기 상세한 설명보다는 후술하는 특허청구범위에 의하여 나타내어지며, 특허청구범위의 의미 및 범위 그리고 그 등가 개념으로부터 도출되는 모든 변경 또는 변형된 형태가 본 발명의 범위에 포함되는 것으로 해석되어야 한다.

【발명의 효과】

<100> 상술한 바와 같이, 본 발명에 의하면 정상인 및 암 질환개체의 혈청 프로테오솜을 이미지 마이닝 기법을 적용하여 질환에 특이적인 스팟에 해당하는 특징을 추출하여 데이터 베이스화하고, 분석 대상 혈청 프로테오솜을 프로테오솜 표본과 비교함으로써 암 질환 발병 여부를 용이하게 확인할 수 있으며, 이에 따라 암 질환의 조기 발견이 가능하게 된다. 또한, 퍼지 룰 기반 분류 방법을 도입함으로써 암 질환 발병 여부를 확인하는 것에 더하여 암 질환의 진행 상태를 알 수 있고 예후를 추적할 수 있어 질환을 가진 개체의 상황에 적합한 의료 처치를 수행할 수 있게 된다.

【특허청구범위】**【청구항 1】**

혈청 프로테옴을 입력받기 위한 입력 수단;

다수의 정상 개체 및 암 질환 개체로부터 채취한 혈청 프로테옴을 상기 입력 수단을 통해 입력받아 2차원 이미지로 변환한 후, 암 질환에 특이적인 특징이 되는 스팟을 추출하고, 상기 추출된 특징 데이터 중에서 최적의 특징 데이터를 식별하여 프로테옴 표본을 생성하고, 분석 대상 혈청 프로테옴을 2차원 이미지로 변환하고 특징을 추출하기 위한 프로테옴 표본 생성 수단;

상기 프로테옴 표본 생성 수단에 의해 추출된 분석 대상 혈청 프로테옴의 패턴을 상기 프로테옴 표본의 패턴과 맵핑시켜 유사 여부를 판단하기 위한 프로테옴 비교 수단;

상기 프로테옴 비교 수단의 비교 결과 분석 대상 혈청 프로테옴의 패턴이 정상 개체의 혈청 프로테옴 패턴과 유사한 경우에는 분석 대상 혈청 프로테옴을 정상으로 판단하고 그렇지 않을 경우에는 암 질환이 발병한 것으로 판단하기 위한 질환 분석 수단; 및

상기 질환 분석 수단의 분석 결과를 출력하기 위한 출력 수단;

을 구비하는 프로테옴 이미지 마이닝을 이용한 암 질환 분석 시스템.

【청구항 2】

제1항에 있어서,

상기 암 질환이 유방암, 난소암, 위암, 간암, 자궁암, 폐암, 대장암, 췌장암 또는 전립선암인 것을 특징으로 하는 암 질환 분석 시스템.

【청구항 3】

제1항에 있어서,

상기 프로테옴 표본이 표1의 스팟들로부터 1 이상 선택되는 것을 특징으로 하는 암 질환 분석 시스템.

【청구항 4】

제1항 내지 제3항의 어느 한 항에 있어서,

상기 프로테옴 표본 생성 수단은 상기 프로테옴 표본을 구성하는 각각의 스팟으로부터 상호 관계성을 추출하고 규칙화함으로써 특징 데이터를 식별하는 프로테옴 이미지 마이닝을 이용한 암 질환 분석 시스템.

【청구항 5】

제1항 내지 제3항의 어느 한 항에 있어서,

상기 질환 분석 수단은 상기 분석 대상 프로테옴이 암 질환개체의 프로테옴인 경우 질환의 진행 정도 및 예후를 예측하고, 상기 분석 대상 프로테옴이 정상 개체의 프로테옴인 경우 암 질환 발병 가능성을 예측하는 프로테옴 이미지 마이닝을 이용한 암 질환 분석 시스템.

【청구항 6】

제1항 내지 제3항의 어느 한 항에 있어서,

상기 암 질환 분석 시스템은 상기 혈청 프로테옴을 제공한 정상 개체 및 암 질환 개체의 정보, 상기 분석 대상 혈청 프로테옴 제공 개체의 정보를 암호화하기 위한 암호화 수단을 더 구비하는 프로테옴 이미지 마이닝을 이용한 암 질환 분석 시스템.

【청구항 7】

제1항 내지 제3항의 어느 한 항에 있어서,
상기 프로테옴 표본 생성 수단은 상기 혈청 프로테옴의 2차원 이미지로부터 의미 있는 특징 데이터를 획득하기 위한 전처리 수단; 및
상기 전처리 수단으로부터 획득한 특징 데이터로부터 분석 대상 혈청 프로테옴의 정상 여부를 확인하기 위한 진화적 분류 수단;
을 더 구비하는 프로테옴 이미지 마이닝을 이용한 암 질환 분석 시스템.

【청구항 8】

제7항에 있어서,
상기 프로테옴 표본 생성 수단은 상기 전처리 수단에서 획득한 특징 데이터로부터 상기 혈청 프로테옴을 구성하는 스팟간의 상호 관계를 추출하고, 추출된 상호 관계를 통계적 기법에 의해 분류하기 위한 퍼지 룰 기반 분류 수단을 더 구비하는 프로테옴 이미지 마이닝을 이용한 암 질환 분석 시스템.

【청구항 9】

제8항에 있어서,
상기 퍼지 룰 기반 분류 수단은 상기 혈청 프로테옴의 2차원 이미지로부터 스팟간의 상호 관계성을 산출하고, 산출된 특성들을 통계적 기법에 의해 분류하며, 퍼지 기법을 이용하여 통계적 불확실성을 정량화하기 위한 데이터 맵핑 수단; 및
상기 데이터 맵핑 수단의 결과를 정리하고 규칙화하여 룰 베이스화하기 위한 룰 기반 분류 수단;

을 구비하는 프로테옴 이미지 마이닝을 이용한 암 질환 분석 시스템.

【청구항 10】

제7항에 있어서,

상기 전처리 수단은 상기 혈청 프로테옴의 2차원 이미지로부터 잡음 제거, 이미지 증폭, 직교 투영, 에지 검출을 수행하기 위한 이미지 처리 수단; 및

상기 영상 처리된 2차원 이미지로부터 스팟 형태의 특징을 추출하고, 추출한 각 특징을 라벨링하기 위한 특징 추출 수단;

을 구비하는 프로테옴 이미지 마이닝을 이용한 암 질환 분석 시스템.

【청구항 11】

제7항에 있어서,

상기 진화적 분류 수단은 상기 전처리 수단에서 추출된 특징 데이터들 중 최적의 특징 데이터들을 식별하기 위한 유전 알고리즘 처리 수단; 및

판단 함수 및 분류 오차율을 이용하여 상기 유전 알고리즘 처리 수단에서 식별해 낸 최적의 특징 데이터들의 적합도를 판단하기 위한 서포트 벡터 머신 적용 수단;

을 구비하는 프로테옴 이미지 마이닝을 이용한 암 질환 분석 시스템.

【청구항 12】

정상 개체 및 암 질환 개체의 혈청 프로테옴을 입력받아 2차원 이미지 처리를 수행하고, 특징 데이터를 추출한 후, 상기 특징 데이터로부터 최적의 특징을 산출하여 프로테옴 표본을 생성하고 데이터베이스화하는 제1 단계;

암 질환의 발병 여부를 확인하고자 하는 분석 대상 혈청 프로테옴을 입력받아 상기 분석 대상 혈청 프로테옴을 2차원 이미지화하고 이로부터 특징 데이터를 추출하는 제2 단계; 및

상기 분석 대상 혈청 프로테옴의 패턴 구조를 상기 프로테옴 표본과 비교하여, 비교 결과에 따라 분석 대상 혈청 프로테옴이 정상인지 암 질환 상태인지 확인하고 분석하는 제3 단계;

를 포함하는 프로테옴 이미지 마이닝을 이용한 암 질환 분석 방법.

【청구항 13】

제12항에 있어서,

상기 암 질환이 유방암, 난소암, 위암, 간암, 자궁암, 폐암, 대장암, 췌장암 또는 전립선암인 것을 특징으로 하는 암 질환 분석 방법.

【청구항 14】

제12항에 있어서,

상기 프로테옴 표본이 표1의 스팟들로부터 1 이상 선택되는 것을 특징으로 하는 암 질환 분석 방법;

【청구항 15】

제12항 내지 제14항의 어느 한 항에 있어서,

상기 제1 단계는 혈청 프로테옴의 2차원 이미지로부터 경험적 지식, 통계적 방식을 적용하여 상기 혈청 프로테옴을 구성하는 스팟간의 상호 관계를 추출하고, 추출된 상호

관계를 통계적 기법에 의해 분류하는 단계를 더 포함하는 프로테옴 이미지 마이닝을 이용한 암 질환 분석 방법.

【청구항 16】

제12항 내지 제14항의 어느 한 항에 있어서,

상기 제3 단계는 상기 분석 대상 혈청 프로테옴의 현재 상태 및 예후를 예측하는 단계를 더 포함하는 프로테옴 이미지 마이닝을 이용한 암 질환 분석 방법.

【청구항 17】

제12항 내지 제14항의 어느 한 항에 있어서,

상기 암 질환 여부를 확인하는 제3 단계는 상기 분석 대상 혈청 프로테옴에 프로테옴 표본 생성 시 추출된 특징들 및 판단 함수를 적용하여 정상 또는 질환 여부를 분류하는 패턴 매칭 과정; 및

상기 2차원 프로테옴 이미지에 포함된 스팟 상호 간의 관계 및 분류 결과를 추론하기 위한 상세 분류 과정;

을 포함하는 프로테옴 이미지 마이닝을 이용한 암 질환 분석 방법.

【청구항 18】

제12항 내지 제14항의 어느 한 항에 있어서,

상기 제3 단계 이후 상기 분석 대상 혈청 프로테옴 및 분석 결과를 데이터베이스화하는 단계를 더 포함하는 프로테옴 이미지 마이닝을 이용한 암 질환 분석 방법.

【청구항 19】

제12항 내지 제14항의 어느 한 항에 있어서,

상기 프로테옴 표본을 생성하는 상기 제1 단계는 상기 2차원 프로테옴 이미지로부터 잡음 제거, 이미지 증폭, 직교 투영, 에지 검출을 수행하는 이미지 처리 단계, 영상 처리된 상기 2차원 프로테옴 이미지로부터 스팟 형태의 기본적인 특징들을 추출하고, 추출한 각 특징들을 라벨링하여 특징 데이터를 생성하는 특징 추출 단계로 이루어지는 전 처리 과정; 및

상기 전처리 과정에서 추출된 특징 데이터들 중 분류에 결정적인 역할을 담당할 최적의 특징들을 식별하기 위한 유전 알고리즘을 수행하고, 서포트 벡터 머신을 적용하여 판단 함수 및 분류 오차율을 이용하여 유전 알고리즘에서 식별해 낸 최적의 특징 데이터들의 적합도를 판단하여 최적의 특징 데이터 및 판단 함수를 추출하는 진화적 분류 과정;

을 포함하는 프로테옴 이미지 마이닝을 이용한 암 질환 분석 방법.

【청구항 20】

제19항에 있어서,

상기 프로테옴 표본을 생성하는 상기 제1 단계는 상기 전처리 과정 이후, 상기 전처리 과정의 결과인 열청 프로테옴의 2차원 이미지로부터 스팟간의 상호 관계성을 산출하고, 산출된 특성들을 통계적 기법에 의해 분류하며, 퍼지 기법을 이용하여 통계적 불확실성을 정량화하는 퍼지 데이터 맵핑 과정;

상기 데이터 맵핑 과정의 결과를 정리하고 규칙화하여 최종적으로 룰 베이스화하는
룰 기반 분류 과정;

을 더 포함하는 프로테옴 이미지 마이닝을 이용한 암 질환 분석 방법.

【청구항 21】

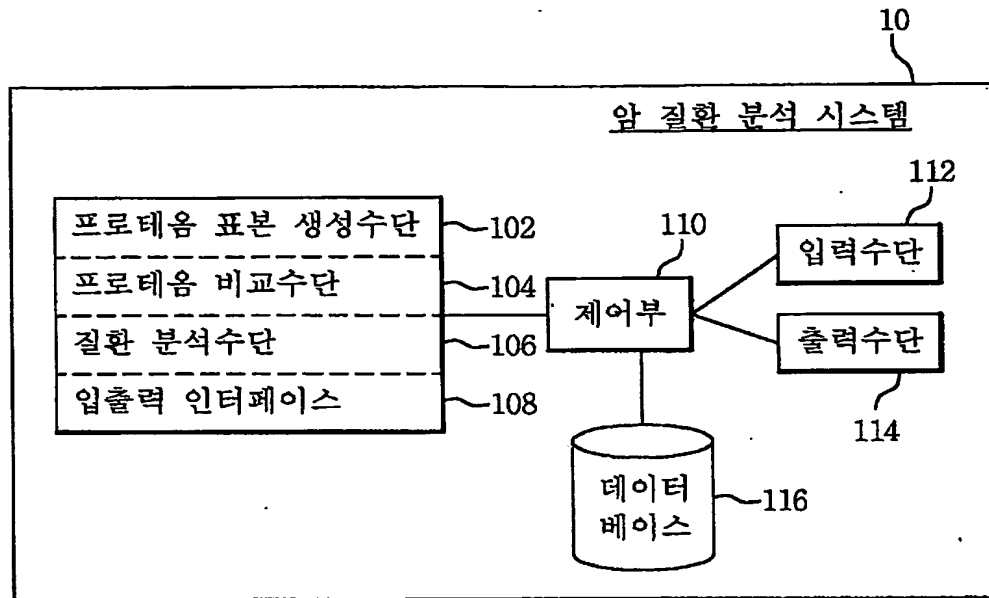
프로테옴 패턴의 분석을 통한 암 질환 진단에 있어서, 상기 프로테옴 패턴이 표1의
스팟들로부터 1이상 선택되는 것을 특징으로 하는 암 진단 마커.

【청구항 22】

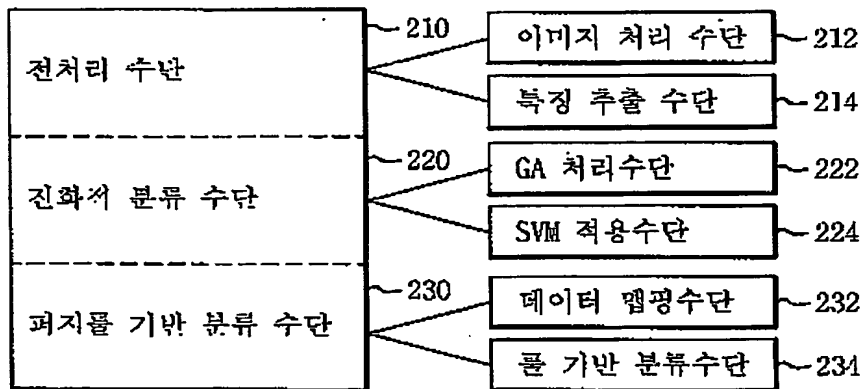
제21에 있어서, 상기 암 질환이 유방암, 난소암, 위암, 간암, 자궁암, 폐암, 대장
암, 췌장암 또는 전립선암인 것을 특징으로 하는 암 진단 마커.

【도면】

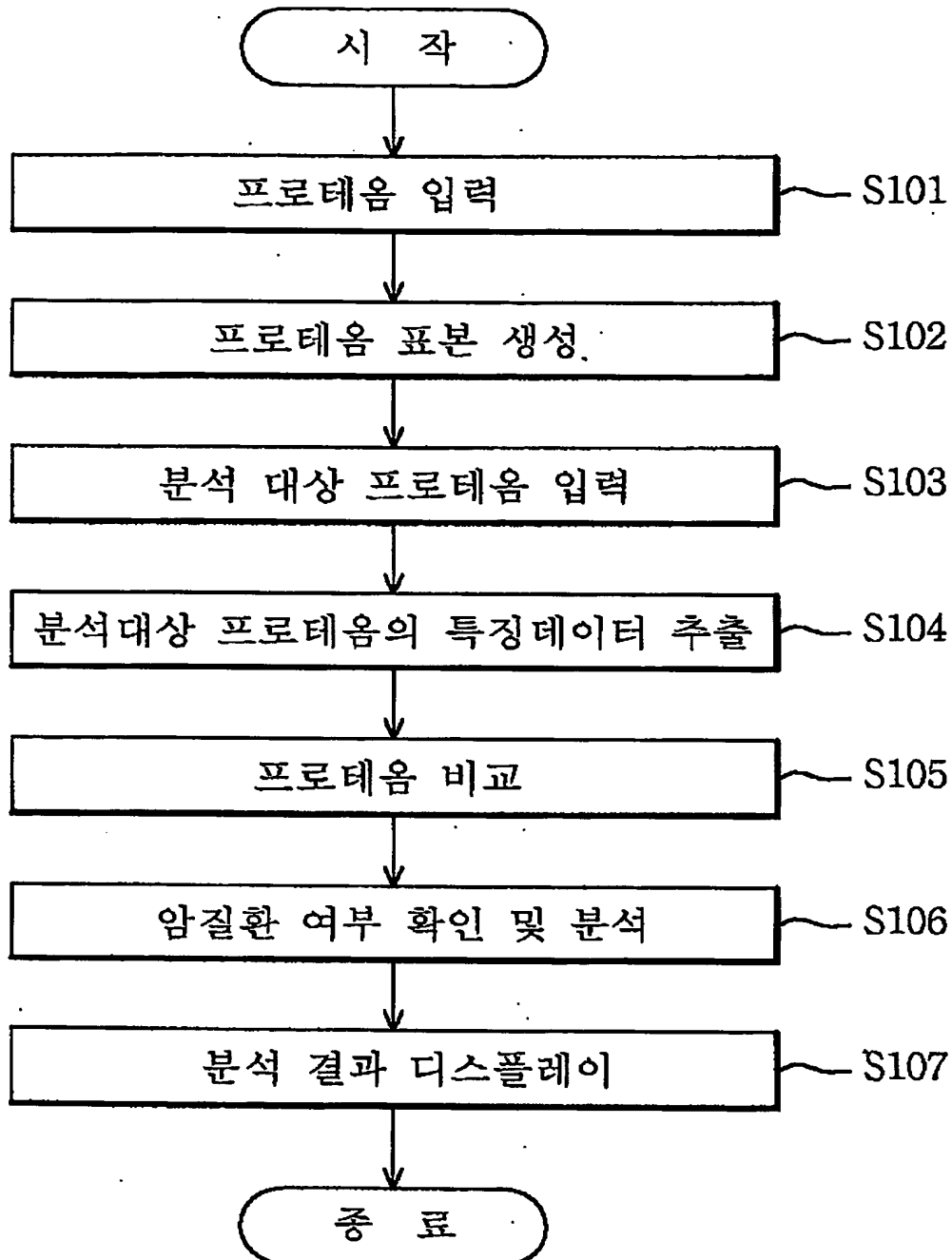
【도 1】



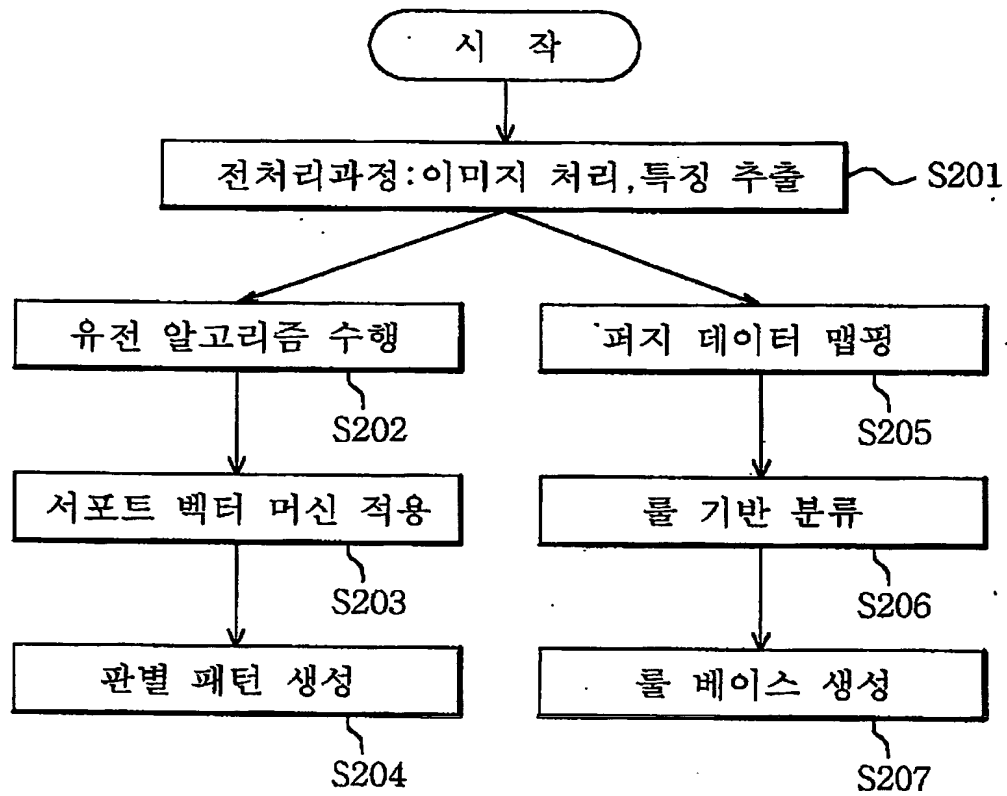
【도 2】



【도 3】

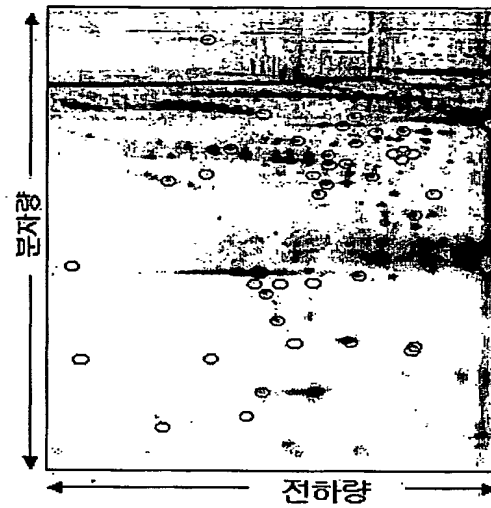


【도 4】

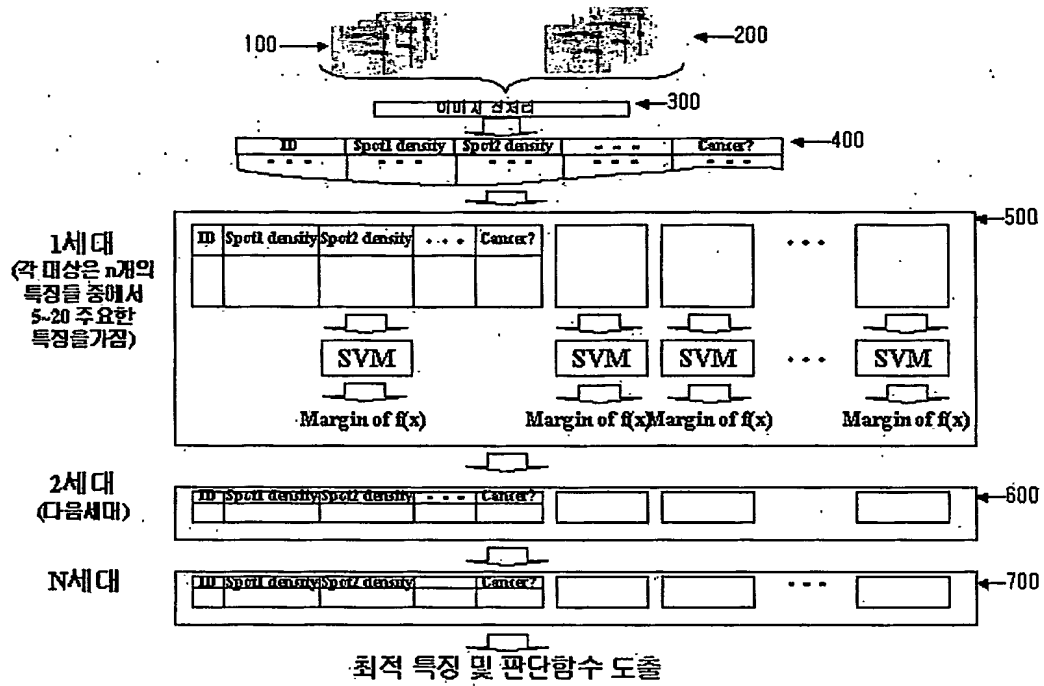


【도 5】

혈청 프로테옴 이미지

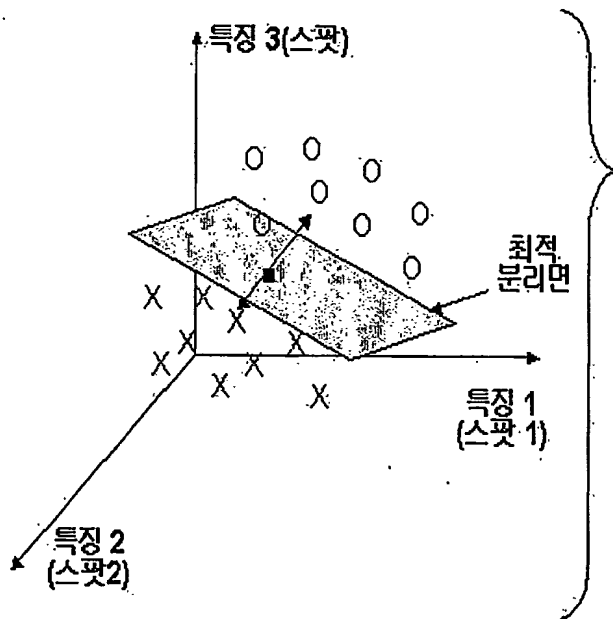


【도 6】



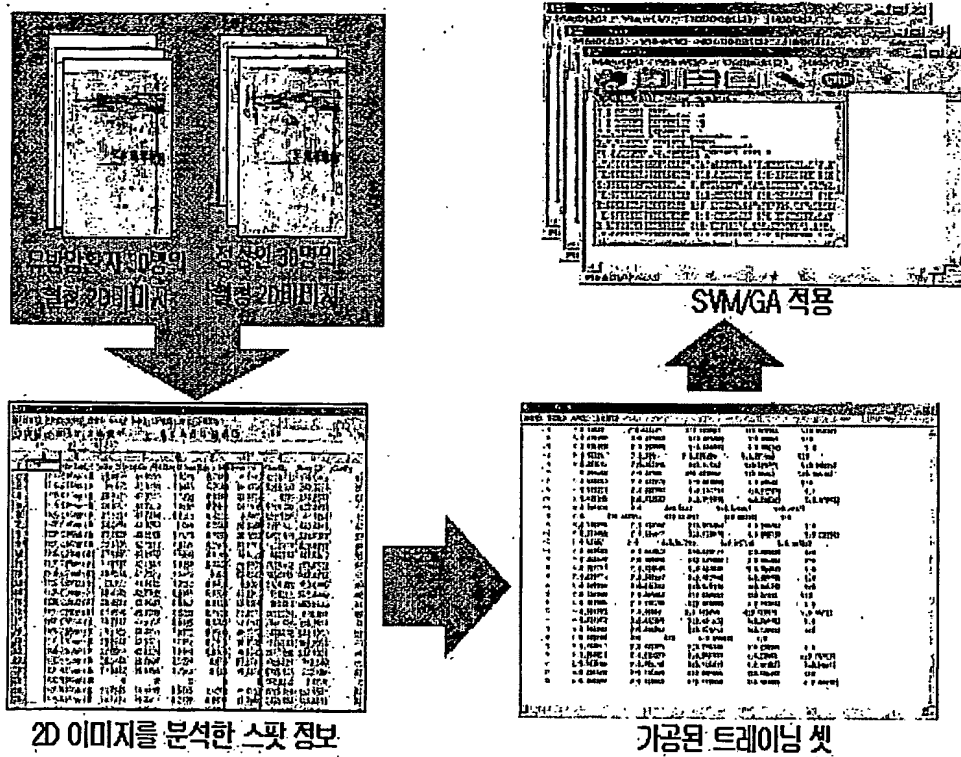
【도 7】

Support Vector Machine



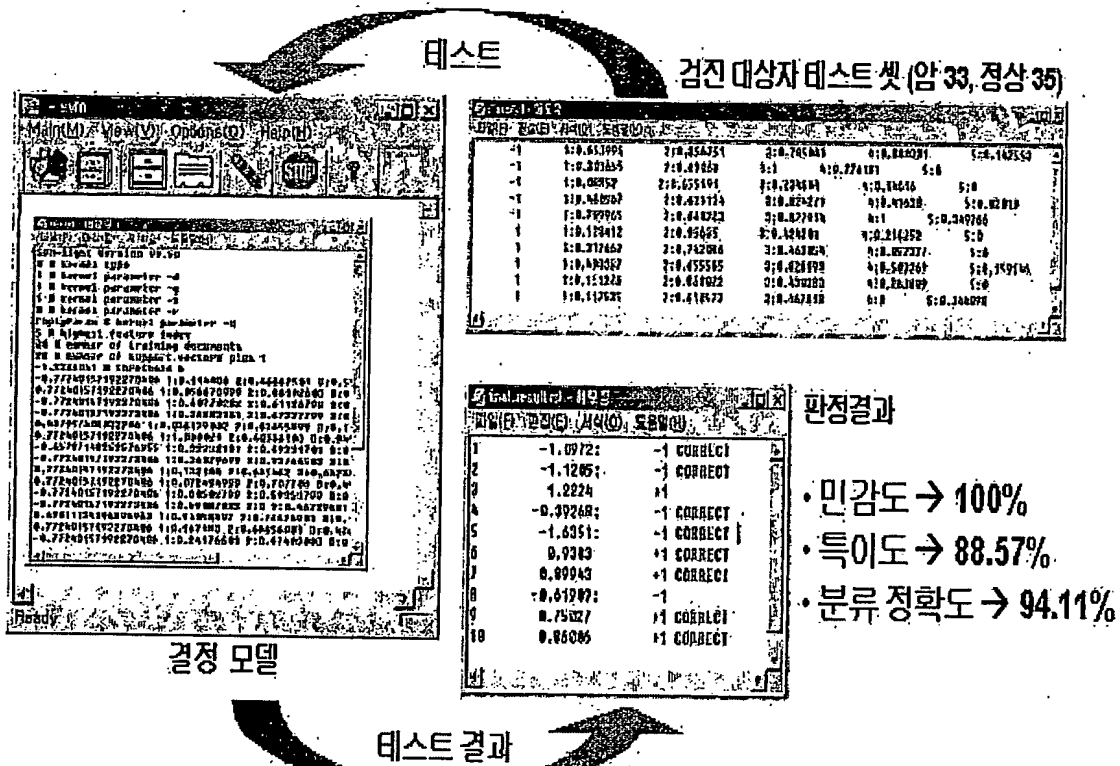
【도 8】

● SVM/GA를 이용한 유방암 진단 - 트레이닝 단계



【도 9】

● SVM/G를 이용한 유방암 진단-테스팅(판별)



【도 10】

34	Value: 0.120890	CORRECT
35	Value: 0.152650	CORRECT
36	Value: 0.140676	CORRECT
37	Value: 0.471308	CORRECT
38	Value: 0.004617	CORRECT
39	Value: 0.721874	CORRECT
40	Value: 0.422242	CORRECT
41	Value: 0.205003	CORRECT
42	Value: 0.026667	CORRECT
43	Value: 0.396419	CORRECT
44	Value: 0.208674	CORRECT
45	Value: 0.370378	CORRECT
46	Value: 0.184488	CORRECT
47	Value: 0.088871	CORRECT
48	Value: 0.250706	CORRECT
49	Value: 0.222000	CORRECT
50	Value: 0.000513	
51	Value: 0.540522	CORRECT
52	Value: 0.711418	CORRECT
53	Value: 0.157877	CORRECT
54	Value: 0.031810	CORRECT
55	Value: 0.030804	CORRECT
56	Value: 0.210711	CORRECT
57	Value: 0.250744	CORRECT
58	Value: 0.124807	CORRECT
59	Value: 0.384526	CORRECT
60	Value: 0.184400	CORRECT
61	Value: -0.107482	
62	Value: 0.052921	CORRECT
63	Value: 0.155490	CORRECT
64	Value: 0.356410	CORRECT
65	Value: 0.025654	CORRECT
66	Value: -0.256264	
67	Value: 0.747711	CORRECT
68	Value: -0.242207	

1	Value: -0.028090	CORRECT
2	Value: -0.002021	CORRECT
3	Value: -0.010359	CORRECT
4	Value: -0.142069	CORRECT
5	Value: 0.402004	CORRECT
6	Value: -0.005008	CORRECT
7	Value: -0.254395	CORRECT
8	Value: -0.549400	CORRECT
9	Value: -0.008596	CORRECT
10	Value: 0.111672	CORRECT
11	Value: -0.247110	CORRECT
12	Value: 0.026007	CORRECT
13	Value: -0.108022	CORRECT
14	Value: -0.444885	CORRECT
15	Value: 0.076969	CORRECT
16	Value: -0.113057	CORRECT
17	Value: 0.356620	CORRECT
18	Value: -0.083508	CORRECT
19	Value: -0.102726	CORRECT
20	Value: 0.002592	CORRECT
21	Value: -0.719689	CORRECT
22	Value: 0.070134	CORRECT
23	Value: -0.070287	CORRECT
24	Value: -0.114127	CORRECT
25	Value: 0.300866	CORRECT
26	Value: -0.077774	CORRECT
27	Value: 0.063312	CORRECT
28	Value: -0.015094	CORRECT
29	Value: -0.001590	CORRECT
30	Value: 0.240725	CORRECT
31	Value: -0.551444	CORRECT
32	Value: 0.101061	CORRECT
33	Value: -0.050204	CORRECT